

Coordinate Update Algorithms for Robust Power Loading for the MU-MISO Downlink with Outage Constraints

Foad Sohrabi and Timothy N. Davidson*

Abstract—We consider the problem of power allocation for the single-cell multi-user (MU) multiple-input single-output (MISO) downlink with quality-of-service (QoS) constraints. The base station acquires an estimate of the channels and, for a given beamforming structure, designs the power allocation so as to minimize the total transmission power required to ensure that target signal-to-interference-and-noise ratios at the receivers are met, subject to a specified outage probability. We consider scenarios in which the errors in the base station's channel estimates can be modelled as being zero-mean and Gaussian. Such a model is particularly suitable for time division duplex (TDD) systems with quasi-static channels, in which the base station estimates the channel during the uplink phase. Under that model, we employ a precise deterministic characterization of the outage probability to transform the chance-constrained formulation to a deterministic one. Although that deterministic formulation is not convex, we develop a coordinate descent algorithm that can be shown to converge to a globally optimal solution when the starting point is feasible. Insight into the structure of the deterministic formulation yields approximations that result in coordinate update algorithms with good performance and significantly lower computational cost. The proposed algorithms provide better performance than existing robust power loading algorithms that are based on tractable conservative approximations, and can even provide better performance than robust precoding algorithms based on such approximations.

Index Terms—Broadcast channel, downlink beamforming, robust power loading, chance constraints, interference functions.

I. INTRODUCTION

It has long been recognized that the provision of multiple antennas at the transmitter of a downlink system has the potential to significantly improve the efficiency with which messages can be communicated from the base station to the receivers; e.g., [1]–[3]. In scenarios in which the base station has perfect knowledge of the state of the channels to each of the single-antenna receivers (and has independent messages to send to them), the dirty paper coding scheme is optimal in the sense it enables the system to achieve

any rate tuple in the capacity region [2], [4]. Since that scheme is quite complicated to implement, a variety of simpler sequential interference pre-subtraction schemes, such as those based on Tomlinson-Harashima precoders [5]–[7] and vector perturbation precoding [8] have been considered. Even simpler schemes based on linear precoding [3], [9]–[11], have also been considered, and we will consider the linear case herein.

In the case of fixed-rate traffic, one approach to the design of the linear precoder is to minimize the power that is required to enable reliable communication to each receiver at their specified target rate. For narrowband systems in which the receivers have a single antenna, that quality-of-service (QoS) problem is equivalent to minimizing the transmission power required to satisfy a signal-to-interference-and-noise (SINR) constraint at each receiver; i.e., min power subject to $\text{SINR}_k \geq \gamma_k$; e.g., [9], [10]. Under the assumption that the transmitter can be provided with accurate channel state information (CSI), without expending a significant fraction of the channel resources, optimal linear precoders for a variety of such quality-of-service (QoS) problems have been obtained; e.g., [9], [10], [12]–[16].

In practice, however, the CSI that can be made available at the transmitter is imperfect, due to estimation errors, quantization, feedback delay, feedback errors, and other effects; e.g., [17], [18]. For the QoS problems that we will consider herein, a straightforward approach to dealing with the resulting uncertainty in the CSI is to perform the design as if the channel estimates were correct, but to increase the SINR targets, γ_k , in order to increase the likelihood that this “mismatched” design meets the original requirements. A more sophisticated approach is to incorporate a model for the uncertainty into the transmitter design. One approach to doing that is to adopt a bounded model for the uncertainty and to design a transmitter that satisfies the QoS requirements even for the worst case of the uncertainties admitted by the model; i.e., min power subject to $\text{SINR}_k \geq \gamma_k$ for all admitted uncertainties; e.g., [19]–[22]. In this paper we consider an alternative approach in which the uncertainty is modelled probabilistically and the QoS requirements are to be satisfied up to a given probability of outage; i.e., min power subject to $\Pr(\text{SINR}_k \geq \gamma_k) \geq 1 - \epsilon_k$; e.g., [23]–[28].

The focus of this paper will be on scenarios in which the uncertainty can be modelled as a zero-mean Gaussian random variable with a given covariance. These scenarios include the case of (single-cell) time-division duplex (TDD) systems operating in quasi-static channels, in which the dominant com-

Manuscript accepted and to appear in IEEE Transactions on Signal Processing, 2016. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada under grant RGPIN-2015-06631. A preliminary exposition of some of the ideas in this work, in the context of a slightly different receiver structure, appears in *Proc. 2013 IEEE Int. Conf. Acoust. Speech, Signal Process.*

F. Sohrabi was with, and T. N. Davidson is with, the Department of Electrical and Computer Engineering, McMaster University, 1280 Main Street West, Hamilton, Ontario, Canada, L8S 4K1. Tel: +1 905 525 9140, Ext. 27352. Fax: +1 905 521 2922. Email: davidson@mcmaster.ca. F. Sohrabi is now with the Department of Electrical and Computer Engineering at the University of Toronto. Email: foad.sohrabi@mail.utoronto.ca.

ponent of the uncertainty arises from the channel estimation error of the (unbiased) estimator on the uplink, and systems in which the channel variation is tracked by the transmitter using variants of the Kalman filtering techniques that have been proposed for receivers; e.g., [29], [30].

One approach to finding good solutions to outage-based QoS problems is to seek a deterministic approximation of the outage constraint that is conservative and can be represented in a form that is convex in design variables. The conservative nature of the approximation means that any point that satisfies the constraints in the resulting restricted optimization problem will satisfy the original outage constraints, and the convex nature of approximations in the approaches means that a globally optimal solution to the restricted optimization problem can be efficiently found. This approach has led to effective techniques for finding good linear precoders [24], [26], [28], and good “power loading” techniques for cases in which the directions of transmission have already been chosen [25], [27].

The proposed approach is somewhat different in that it does not involve an approximation of the outage constraint, but employs a precise deterministic representation for the case of Gaussian uncertainties [31], [32]. Unlike the previous approaches, the resulting optimization problem is not convex, but we develop a straightforward cyclic coordinate descent algorithm that, through connections with the framework of interference functions [33], can be shown to converge to an optimal solution when the starting point satisfies the outage constraints. (A related fixed-point algorithm was developed concurrently in [34].) Even when terminated quite early, this algorithm typically provides superior performance to that of the globally optimal solutions to the conservative approximation.

Insight developed from our initial implementation of the basic principle is then used to construct a more computationally efficient power-loading technique for the case of nominally “zero-forcing” beamforming directions. While that technique does involve a conservative approximation, the structure of the approximation is quite different from those that have been previously applied, and numerical experience suggests that it can be significantly less conservative. Interestingly, in some important scenarios the lower level of conservatism in the approximation means that the proposed power loading algorithm with nominally zero-forcing directions yields better performance than existing techniques in which the power loading and directions are designed jointly.

The remainder of the paper is organized as follows: The system model for the downlink and the uncertainty model that we will consider are described in Section II. Having established those models, in Sections III-A and III-B, we formally define the problems of robust precoding and robust power loading, respectively. We also review some of the existing approaches to those problems. In Section IV we provide the result from [32] that enables us to write closed-form deterministic expressions for the outage probability (in the case of Gaussian uncertainties). In Section V we present a coordinate descent algorithm for optimal power loading in the case of a generic selection of the beamforming directions. In Section VI we present several tailored algorithms for the

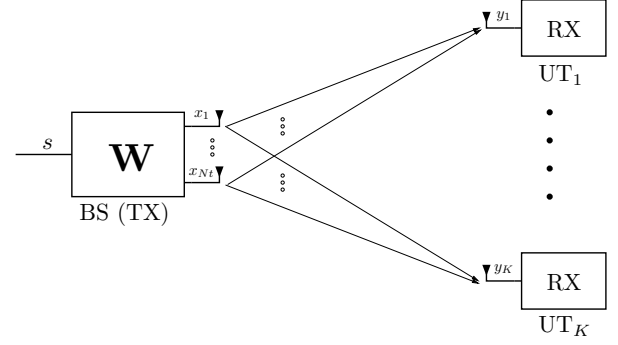


Fig. 1. A single-cell downlink setting with N_t antennas at the base station (BS) and K users, each with a single antenna.

case of the zero-forcing directions (for the base station’s estimates of the channels). These algorithms are based on an approximation of the integrand in the generic method. That approximation enables the integral to be computed using residue techniques, which significantly reduces the computational cost of the algorithm. In Section VII we compare the performance of the proposed algorithms to a number of existing approaches, and conclusions are drawn in Section VIII. Some of the details of the technical results are discussed in the Appendices.

II. SYSTEM MODEL

We consider a narrowband single-cell downlink scenario in which a base station with N_t antennas sends independent messages to K users (unicast transmission), each of which is equipped with a single antenna, as illustrated in Fig. 1. The base station employs linear precoding and the transmitted signal at each channel use is

$$\mathbf{x} = \sum_{k=1}^K \mathbf{w}_k s_k = \mathbf{W} \mathbf{s}, \quad (1)$$

where $\mathbf{w}_k \in \mathbb{C}^{N_t}$ is the beamforming vector for the k^{th} user and forms the k^{th} column of the precoding matrix $\mathbf{W} \in \mathbb{C}^{N_t \times K}$, and s_k is the symbol to be sent to the k^{th} user. We normalize the symbols so that they have unit energy, and since the messages are assumed to be independent we have $E\{\mathbf{s}\mathbf{s}^H\} = \mathbf{I}$. For user k , the received signal can be modelled as

$$y_k = \mathbf{h}_k^H \mathbf{x} + z_k, \quad (2)$$

where $\mathbf{h}_k^H \in \mathbb{C}^{N_t}$ is the row vector of complex channel gains from the transmitting antennas to the k^{th} receiver, and z_k denotes additive noise, which is assumed to be circular complex Gaussian with zero mean and variance σ_k^2 . The received signal can be rewritten as

$$y_k = \mathbf{h}_k^H \mathbf{w}_k s_k + \mathbf{h}_k^H \bar{\mathbf{W}}_k \mathbf{s} + z_k, \quad (3)$$

where $\bar{\mathbf{W}}_k = [\mathbf{w}_1, \dots, \mathbf{w}_{k-1}, \mathbf{0}, \mathbf{w}_{k+1}, \dots, \mathbf{w}_K]$. The first term in (3) is the useful signal term for coherent detection of the message sent to user k , while the second term represents the interference due to the transmissions to the other users and the third term is the noise.

The desire to provide fixed-rate services to several classes of users has led to development of design techniques which

guarantee that a certain quality-of-service (QoS) constraint is satisfied for each user. In the case of coherent single-user detection, in which interference is treated as noise, the QoS is typically specified in terms of the SINR at the k^{th} user,

$$\text{SINR}_k(\mathbf{W}) = \frac{|\mathbf{h}_k^H \mathbf{w}_k|^2}{\mathbf{h}_k^H \bar{\mathbf{W}}_k \bar{\mathbf{W}}_k^H \mathbf{h}_k + \sigma_k^2}. \quad (4)$$

If the BS has perfect knowledge of the channels, $\mathbf{h}_1^H, \dots, \mathbf{h}_K^H$, then for any choice of the beamforming matrix \mathbf{W} , it can compute each receiver's SINR and hence it can adapt its transmission. In a standard scenario, the operator may wish to have the BS adapt its transmission so that it minimizes the power, $E\{\mathbf{x}^H \mathbf{x}\} = \sum_k \|\mathbf{w}_k\|^2 = \text{Tr}(\mathbf{W}\mathbf{W}^H)$, required to provide selected users with a specified target SINR, or declare that it is not possible to meet that specification. That is, the BS may be required to find a set of beamformers $\{\mathbf{w}_k\}$ that solves

$$\min_{\{\mathbf{w}_k \in \mathbb{C}^{N_t}\}_{k=1}^K} \text{Tr}(\mathbf{W}\mathbf{W}^H) \quad (5a)$$

$$\text{subject to } \text{SINR}_k(\mathbf{W}) \geq \gamma_k, \quad \forall k = 1, \dots, K, \quad (5b)$$

or show that no such set of beamformers exists. Here, γ_k is the specified SINR for user k . This problem can be efficiently solved by transforming it into a convex second order cone program [13]; by applying a fixed-point mapping [9]; by applying the (related) notions of ‘uplink-downlink’ or Lagrangian duality [12], [35]; or by relaxing it to a convex semidefinite program and showing that the relaxation is tight [10]. A ‘dedicated’ training phase is then employed so that the k^{th} receiver can estimate the scalar $\mathbf{h}_k^H \mathbf{w}_k$ with an accuracy that is sufficient to perform the coherent detection that is implicit in (4).

In practice, the CSI that can be made available at the BS is imperfect, due to estimation errors, quantization, delays, and other effects; e.g., [17], [18], and yet the design of the downlink precoder with QoS targets remains an important problem. As simple strategy is to perform a ‘mismatched’ design in which the BS’s estimates of the channels are treated as if they were accurate, and the SINR targets γ_k are increased in an attempt to increase the probability that the original targets are satisfied for the actual channels. An alternative strategy is to construct a model for the uncertainty in the BS’s knowledge of the channel and to take that uncertainty into account in the formulation of the design problem. The model for the uncertainty is dependent on the method by which the BS obtains information about the state of the channel.

In this paper, we will consider scenarios in which a Gaussian model for that uncertainty is appropriate. Among several scenarios in which such a model is appropriate, a prominent one is the case of (single-cell) time-division duplex (TDD) systems, in which the base station obtains an estimate of each channel via training and linear estimation on the uplink (e.g., [36]), using the principle of reciprocity of the channel [37]. In that setting, if the channel changes sufficiently slowly (with respect to the ‘ping-pong time’ of the TDD system), and if appropriate RF calibration is performed (e.g., [38]), then the estimation error is the dominant source of the mismatch between the actual and estimated channels. When the linear

estimator is unbiased, the base station’s estimate of the (base-band equivalent) channels can be written as

$$\hat{\mathbf{h}}_k^H = \mathbf{h}_k^H + \mathbf{e}_k^H, \quad (6)$$

where \mathbf{e}_k is an zero-mean Gaussian random variable that is independent of \mathbf{h}_k and has a covariance matrix, \mathbf{C}_k . That is, \mathbf{e}_k is Gaussian with $E\{\mathbf{e}_k\} = \mathbf{0}$ and $E\{\mathbf{e}_k \mathbf{e}_j^H\} = \delta[k-j] \mathbf{C}_k$, where $\delta[\cdot]$ denotes the Kronecker delta. In particular, if the noise on the uplink is circular complex Gaussian with zero mean and covariance $\sigma_{\text{BS}}^2 \mathbf{I}$, then if we treat the channel vector as being deterministic, the best linear unbiased estimator (BLUE) is actually the least-squares estimator and the covariance matrices \mathbf{C}_k depend on the structure and SNR of the training sequence; e.g., [36], [39]. In this setting, the least-squares estimator is also a minimum variance unbiased estimator. In the case of orthogonal training, which is optimal in this setting [36], the covariance matrices reduce to $\mathbf{C}_k = \sigma_e^2 \mathbf{I}$, where σ_e^2 is a function of the training SNR and the length of the training sequence.

III. DESIGN OF OUTAGE-BASED ROBUST DOWNLINK TRANSMITTERS

The focus of this paper is on robust power loading schemes for systems in which the beamforming directions $\mathbf{w}_k / \|\mathbf{w}_k\|$ have been chosen; see Section III-B. However, for completeness we first consider the case of full precoder design.

A. Chance-constrained Robust Precoding

For given SINR targets γ_k , the quality of service constraint is that the probability that $\text{SINR}_k(\mathbf{W}) \geq \gamma_k$ should be greater than $1 - \epsilon_k$, for a pre-specified ‘probability of outage’ ϵ_k . Therefore, given the uncertainty model $\mathbf{h}_k = \hat{\mathbf{h}}_k + \mathbf{e}_k$ and a distribution for \mathbf{e}_k , the problem of interest can be written as

$$\begin{aligned} \min_{\{\mathbf{w}_k \in \mathbb{C}^{N_t}\}_{k=1}^K} & \text{Tr}(\mathbf{W}\mathbf{W}^H) \\ \text{subject to } & \Pr_{\mathbf{e}_k}(\text{SINR}_k(\mathbf{W}) \geq \gamma_k) \geq 1 - \epsilon_k, \quad \forall k, \end{aligned} \quad (7a)$$

$$(7b)$$

where the SINR at user k was defined in (4). The presence of the chance constraints in (7b) makes the problem difficult to tackle directly; the SINR in (4) is the ratio of quadratic functions of the design variables. One approach is to apply a conservative transformation to the SINR constraint in (7b) to convert the problem in (7) into a chance-constrained second-order cone program (SOCP) [26]. By applying various conservative approximations of chance-constrained SOCPs, efficiently-solvable deterministic convex optimization problems are obtained (some are SOCPs, others are semidefinite programs, SDPs). The conservative nature of the approximations means that when these convex problems are feasible,¹ the solution is guaranteed to satisfy the chance constraints in (7b).

The related approach of [28] first applies a semidefinite relaxation to the problem in (7), which yields a semidefinite program (SDP) with chance constraints on quadratic

¹in the sense that there exists at least one choice for the set of beamformers $\{\mathbf{w}_k\}_{k=1}^K$ that satisfies all the constraints in the conservative approximation

functions of a vector of variables. These chance constraints are then conservatively approximated by deterministic convex constraints leading to an SDP formulation. The solution to that SDP formulation is guaranteed to satisfy the original chance constraints in (7b) whenever each of the solution matrices has rank one. Numerical experiments suggest that this is almost always the case, especially when the conservative approximation of the chance constraint results in a spherical uncertainty region, as distinct from the more general elliptical case [28], [40]. Under the zero-mean Gaussian uncertainty model in (6), one of the SDP problem formulations obtained using this approach involves optimizing over Hermitian matrices \mathbf{U}_k that represent $\mathbf{w}_k \mathbf{w}_k^H$, but are not required to be rank 1, and conservatively approximating the chance constraint by a linear matrix inequality [28]. The resulting SDP is

$$\begin{aligned} \min_{\{\mathbf{U}_k \in \mathbb{H}^{N_t \times N_t}\}, \{t_k \geq 0\}} \quad & \sum_k \text{Tr}(\mathbf{U}_k) \\ \text{subject to} \quad & \begin{bmatrix} \mathbf{Q}_k + t_k \mathbf{I} & \mathbf{r}_k \\ \mathbf{r}_k^H & v_k - t_k d_k^2 \end{bmatrix} \succeq \mathbf{0}, \quad \forall k, \end{aligned} \quad (8a)$$

(8b)

where $\mathbf{Q}_k = \mathbf{C}_k^{1/2} \left(\frac{1}{\gamma_k} \mathbf{U}_k - \sum_{j \neq k} \mathbf{U}_j \right) \mathbf{C}_k^{1/2}$, $\mathbf{r}_k = \mathbf{C}_k^{1/2} \left(\frac{1}{\gamma_k} \mathbf{U}_k - \sum_{j \neq k} \mathbf{U}_j \right) \hat{\mathbf{h}}_k$, $v_k = \hat{\mathbf{h}}_k^H \left(\frac{1}{\gamma_k} \mathbf{U}_k - \sum_{j \neq k} \mathbf{U}_j \right) \hat{\mathbf{h}}_k - \sigma_k^2$ and $d_k = \sqrt{\phi^{-1}_{X_{2N_t}^2}(1 - \epsilon_k)/2}$, where $\phi^{-1}_{X_{2N_t}^2}(\cdot)$ is the inverse cumulative distribution function of central Chi-square random variable with $2N_t$ degrees of freedom. We will use the formulation in (8) as a benchmark in the evaluation of the proposed designs.

B. Chance-constrained Robust Power Loading

In robust precoding the directions of transmission, $\check{\mathbf{w}}_k = \mathbf{w}_k / \|\mathbf{w}_k\|_2$, and the power allocated to each direction, $\check{p}_k = \|\mathbf{w}_k\|_2^2$, are found jointly. A potentially simpler approach is to choose the directions $\check{\mathbf{w}}_k$ based on the transmitters' channel estimates $\hat{\mathbf{h}}_k$ and then to seek solutions to the problem in (7) over the K powers, \check{p}_k . It is often more convenient to remove the restriction that the directions be specified in a normalized form, and simply pre-specify vectors. We will pre-specify the directions as not necessarily normalized vectors \mathbf{b}_k and seek a power allocation $\{p_k\}_{k=1}^K$ for these vectors so that $\mathbf{w}_k = \sqrt{p_k} \mathbf{b}_k$. In that case, the total power transmitted is $\sum_{k=1}^K p_k \|\mathbf{b}_k\|_2^2$. If we define $\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_K]$ and the diagonal matrix $\mathbf{P} = \text{Diag}(p_1, p_2, \dots, p_K)$, the robust power loading problem can be written as

$$\min_{\{p_k \geq 0\}} \text{Tr}(\mathbf{B} \mathbf{P} \mathbf{B}^H) \quad (9a)$$

$$\text{s.t. } \Pr_{\mathbf{e}_k} \left(\frac{(\hat{\mathbf{h}}_k^H + \mathbf{e}_k^H) \mathbf{b}_k \mathbf{b}_k^H (\hat{\mathbf{h}}_k + \mathbf{e}_k)}{(\hat{\mathbf{h}}_k^H + \mathbf{e}_k^H) \bar{\mathbf{B}}_k \mathbf{P} \bar{\mathbf{B}}_k^H (\hat{\mathbf{h}}_k + \mathbf{e}_k) + \sigma_k^2} \geq \gamma_k \right) \geq 1 - \epsilon_k, \quad \forall k, \quad (9b)$$

where $\bar{\mathbf{B}}_k = [\mathbf{b}_1, \dots, \mathbf{b}_{k-1}, \mathbf{0}, \mathbf{b}_{k+1}, \dots, \mathbf{b}_K]$. A common choice for the precoding matrix \mathbf{B} is the regularized channel inversion precoder for the estimated channel [41]: Given

matrix of channel estimates $\hat{\mathbf{H}} = [\hat{\mathbf{h}}_1, \hat{\mathbf{h}}_2, \dots, \hat{\mathbf{h}}_K]^H$ and a non-negative real number α ,

$$\mathbf{B}_{\text{RCI}} = \hat{\mathbf{H}}^H (\hat{\mathbf{H}} \hat{\mathbf{H}}^H + \alpha \mathbf{I}_K)^{-1}. \quad (10)$$

In the special case when $\alpha = 0$ (and $K \leq N_t$), the nominal zero-forcing precoder is obtained,

$$\mathbf{B}_{\text{ZF}} = \hat{\mathbf{H}}^H (\hat{\mathbf{H}} \hat{\mathbf{H}}^H)^{-1}. \quad (11)$$

In the development of approaches to solve the problem in (9), we will rewrite the chance constraints in (9b) in the form of chance constraints on a quadratic function of a standard complex Gaussian random variable, $\delta_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$, namely,

$$\Pr_{\delta_k} (\delta_k^H \mathbf{Q}_k \delta_k + 2 \text{Re}(\delta_k^H \mathbf{r}_k) + v_k \geq 0) \geq 1 - \epsilon_k, \quad (12)$$

where $\mathbf{Q}_k = \mathbf{C}_k^{1/2} \left(\frac{p_k}{\gamma_k} \mathbf{b}_k \mathbf{b}_k^H - \bar{\mathbf{B}}_k \mathbf{P} \bar{\mathbf{B}}_k^H \right) \mathbf{C}_k^{1/2}$, $\mathbf{r}_k = \mathbf{C}_k^{1/2} \left(\frac{p_k}{\gamma_k} \mathbf{b}_k \mathbf{b}_k^H - \bar{\mathbf{B}}_k \mathbf{P} \bar{\mathbf{B}}_k^H \right) \hat{\mathbf{h}}_k$ and $v_k = \hat{\mathbf{h}}_k^H \left(\frac{p_k}{\gamma_k} \mathbf{b}_k \mathbf{b}_k^H - \bar{\mathbf{B}}_k \mathbf{P} \bar{\mathbf{B}}_k^H \right) \hat{\mathbf{h}}_k - \sigma_k^2$. By writing the chance constraints in this form, a number of existing conservative deterministic approximations to the chance constraint can be applied in a straightforward way; see [28]. For example, given \mathbf{B} , the solution to the following SDP yields powers $\{p_k\}$ that satisfy the constraints in (9),

$$\min_{\{p_k \geq 0\}, \{t_k \geq 0\}} \text{Tr}(\mathbf{B} \mathbf{P} \mathbf{B}^H) \quad (13a)$$

$$\text{subject to} \quad \begin{bmatrix} \mathbf{Q}_k + t_k \mathbf{I} & \mathbf{r}_k \\ \mathbf{r}_k^H & v_k - t_k d_k^2 \end{bmatrix} \succeq \mathbf{0}, \quad \forall k, \quad (13b)$$

where $d_k = \sqrt{\phi^{-1}_{X_{2N_t}^2}(1 - \epsilon_k)/2}$. That said, since (13) is based on a conservative approximation, the absence of a solution to (13) does not necessarily mean that there are no powers that satisfy (9).

The goal of this paper is to propose robust power loading algorithms that reduce conservatism and may reduce the computational cost. Rather than being based on seeking tractable convex, but conservative formulations, the proposed approach is based on a closed-form expression for the probability that the SINR constraint is satisfied.

IV. A CLOSED-FORM EXPRESSION FOR THE CDF OF A QUADRATIC FUNCTION OF A GAUSSIAN RANDOM VECTOR

The proposed approaches to the robust power loading problem in (9) will be based on the following closed-form expression for the cumulative distribution function (CDF) of a quadratic function of a standard circular complex Gaussian random vector [32]. To state that expression, we will use the notation \mathbf{i} to denote $\sqrt{-1}$, and $\|\mathbf{u}\|_{\mathbf{M}}^2$ to denote $\mathbf{u}^H \mathbf{M} \mathbf{u}$.

Lemma 1 ([32]): Given a deterministic Hermitian symmetric matrix \mathbf{M} and a deterministic vector \mathbf{z} , for the standard circular complex Gaussian random vector $\mathbf{x} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$

$$\begin{aligned} \Pr(\|\mathbf{x} - \mathbf{z}\|_{\mathbf{M}}^2 \leq \tau) \\ = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{\tau(\mathbf{i}\omega + \beta)}}{\mathbf{i}\omega + \beta} \frac{e^{-c}}{\det(\mathbf{I} + (\mathbf{i}\omega + \beta)\mathbf{M})} d\omega, \end{aligned} \quad (14)$$

for some $\beta > 0$ such that $\mathbf{I} + \beta \mathbf{M}$ is positive definite. If we let $\mathbf{M} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^H$ denote the eigen decomposition of \mathbf{M} , with λ_m

denoting the eigenvalues arranged in descending order ($\mathbf{\Lambda} = \text{Diag}(\lambda_1, \lambda_2, \dots)$), and if we define $\tilde{\mathbf{z}} = \mathbf{V}^H \mathbf{z}$, the constant c can be written as $c = \sum_{m=1}^M \frac{|\tilde{z}_m|^2 (i\omega + \beta) \lambda_m}{1 + (i\omega + \beta) \lambda_m}$.

The second statement in Lemma 1 is slightly more general than that in [32] because it does not require \mathbf{M} to be invertible. (A complete proof is provided in [42].) In the application herein, that is important when there are fewer active users than transmitting antennas.

As an aside, we observe that alternative approaches to those proposed below can be developed by considering circular complex Gaussian random variables $\mathbf{g}_k \sim \mathcal{CN}(\hat{\mathbf{h}}_k, \mathbf{C}_k)$, rewriting the probability on the left hand side of (9b) as

$$\Pr_{\mathbf{g}_k}(\mathbf{g}_k^H (\frac{p_k}{\gamma_k} \mathbf{b}_k \mathbf{b}_k - \bar{\mathbf{B}}_k \mathbf{P} \bar{\mathbf{B}}_k^H) \mathbf{g}_k - \sigma_k^2 \geq 0) \geq 1 - \epsilon_k, \quad (15)$$

and employing the closed-form expression for this probability that can be obtained by applying the residue-based analysis in [31]. While that would be effective for the generally-applicable algorithm developed in Section V, the infinite series in the expression obscure insight. The insight that lead to the development of the tailored algorithms in Section VI arose from the expression in Lemma 1.

V. FEASIBLE COORDINATE DESCENT ALGORITHM

By reformatting the chance constraint in (12) as

$$\Pr(\|\delta_k - \mathbf{a}_k\|_{(-\mathbf{Q}_k)}^2 \leq \tau_k) \geq 1 - \epsilon_k, \quad (16)$$

where $\mathbf{a}_k = -\mathbf{C}_k^{-1/2} \hat{\mathbf{h}}_k$ and $\tau_k = v_k - \mathbf{a}_k^H \mathbf{Q}_k \mathbf{a}_k$, we can employ Lemma 1 and rewrite the robust power loading problem in (9) in a form that is no longer chance-constrained, but is deterministically constrained:

$$\begin{aligned} \min_{\{p_k \geq 0\}} \quad & \text{Tr}(\mathbf{B} \mathbf{P} \mathbf{B}^H) \\ \text{s.t.} \quad & \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{\tau_k(i\omega + \beta)}}{i\omega + \beta} \frac{e^{-c_k}}{\det(\mathbf{I} - (i\omega + \beta) \mathbf{Q}_k)} d\omega \\ & \geq 1 - \epsilon_k, \quad \forall k, \end{aligned} \quad (17a) \quad (17b)$$

where \mathbf{Q}_k and τ_k were defined following (12) and (16), respectively, and c_k has a format analogous to the format of c in Lemma 1. This deterministic problem is not convex. However, since the integral in (17b) is equivalent to $\Pr(\text{SINR}_k \geq \gamma_k)$, we can interpret the behaviour of the integral by looking at the definition of the SINR. For fixed-direction beamformers, $\{\mathbf{b}_k\}_{k=1}^K$, we can rewrite the SINR in (4) in terms of the powers as

$$\text{SINR}_k = \frac{|\mathbf{h}_k^H \mathbf{b}_k|^2 p_k}{\sum_{j \neq k} |\mathbf{h}_k^H \mathbf{b}_j|^2 p_j + \sigma_k^2}. \quad (18)$$

We can also rewrite the term $\bar{\mathbf{B}}_k \mathbf{P} \bar{\mathbf{B}}_k^H$ that appears in (9b) and (12) as $\sum_{j \neq k} p_j \mathbf{b}_j \mathbf{b}_j^H$. Using insight from (18) and (9b), by making the above substitution in (12), it can be seen that for fixed p_j , $j \neq k$, the integral in (17b) is increasing in p_k , and that for fixed p_k , the integral in (17b) is decreasing in each p_j , $j \neq k$. This observation suggests the development of a cyclic coordinate descent algorithm (e.g., [43, Sec. 2.7]) in which we start from a power allocation $\{p_k\}_{k=1}^K$ that forms a feasible

point² for the problem in (17), and at the k^{th} step of the i^{th} cycle we seek to reduce the value of p_k given the current values of the other powers, while maintaining feasibility. A feature of that approach is that due to the above-mentioned features of (12), at the k^{th} step of the i^{th} cycle we need only consider the k^{th} constraint in (17b); decreasing p_k will not violate any of the other constraints. Furthermore, in concurrent work [34] it was shown that the power allocation problem in (9), and hence the equivalent problem in (17), can be viewed in the framework of standard interference functions [33]. As a result, if the initial power allocation vector is feasible, and if we solve for the minimum feasible p_k at each step (in which case all the constraints in (17b) are satisfied with equality), the cyclic coordinate descent algorithm described above will converge a globally optimal solution [14], [33]. The particular algorithm that we will develop below involves some additional parameters that will facilitate computational tradeoffs in the algorithm. Although we will describe the algorithm with each step in the cycle in the natural order, the principles apply to other orderings and even certain asynchronous updating schemes [33].

To put these principles into practice, we let $\mathbf{P}^{(0)}$ denote the diagonal matrix containing the initial feasible power allocation. (We will discuss techniques for finding such a $\mathbf{P}^{(0)}$ below.) At the k^{th} step of the i^{th} cycle we choose a value for $\Delta_k^{(i)} \leq \Delta_k^{(i-1)}$ and perform a bisection search on the interval $[0, p_k^{(i-1)}]$ for a value of p_k such that the probability that $\text{SINR}_k \geq \gamma_k$ lies in the interval $[1 - \epsilon_k, 1 - \epsilon_k + \Delta_k^{(i)}]$; i.e.,

$$1 - \epsilon_k \leq \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{\tau_k(i\omega + \beta)}}{i\omega + \beta} \frac{e^{-c_k}}{\det(\mathbf{I} - (i\omega + \beta) \mathbf{Q}_k)} d\omega \leq 1 - \epsilon_k + \Delta_k^{(i)} \quad (19)$$

with $c_k^{(i)}$, $\tau_k^{(i)}$, $\mathbf{Q}_k^{(i)}$ being calculated using $p_1^{(i)}, \dots, p_{k-1}^{(i)}, p_{k+1}^{(i)}, \dots, p_K^{(i-1)}$ and the midpoint for the current interval in the bisection search for $p_k^{(i)}$.

The algorithm is terminated once we find a power allocation $\mathbf{P}^{(i)}$ such that for each k the probability that $\text{SINR}_k \geq \gamma_k$ lies in the interval $[1 - \epsilon_k, 1 - \epsilon_k + \Delta_k^{\min}]$, where Δ_k^{\min} is a pre-specified bound. (As is implicit in (19), the cycle at which the algorithm terminates is the first one for which $\Delta_k^{(i)} \leq \Delta_k^{\min}$ for all k .) A feature of this algorithm is that at each step in each cycle the power allocation $\{p_1^{(i)}, p_2^{(i)}, \dots, p_k^{(i)}, p_{k+1}^{(i-1)}, \dots, p_K^{(i-1)}\}$ is feasible and hence whenever the algorithm is terminated, the current power allocation will satisfy the specified QoS constraints of the original problem. Furthermore, at each step in each cycle, the objective value decreases, or remains the same. The latter case occurs when the termination criteria is satisfied prior to performing the current coordinate descent step and therefore the algorithm will go on to the next step without any changes in the power allocation.

The parameter Δ_k^{\min} enables us to make tradeoffs between the performance and the computational cost of the algorithm.

²That is, a power allocation $\{p_k\}_{k=1}^K$ for which all K constraints in (17b) are satisfied.

A smaller value for Δ_k^{\min} results in less conservative solutions that are achieved using less transmitted power, but will typically lead to a larger value for the total number of bisection steps (cf. (19)) required for the algorithm to terminate. While the choice of Δ_k^{\min} controls the performance of the algorithm and has the dominant influence on its computational cost, the flexibility to choose $\Delta_k^{(i)}$ at each cycle offers the opportunity to trade the number of bisection steps required in each cycle against the number of cycles required. A variety of intuitively motivated decreasing sequences for $\{\Delta_k^{(i)}\}_i$ could be considered, but our numerical experience suggests that the simple choice of $\Delta_k^{(i)} = \Delta_k^{\min}$, which results in a solution being obtained in a single cycle, leads to an effective implementation.

To complete the description of the algorithm, we need to establish a method to determine a feasible starting point, $\mathbf{P}^{(0)}$. As the feasible set³ in (17) is not necessarily convex, determining whether or not an instance of the problem in (17) is feasible can be computationally demanding task; e.g., [44, Section 5.1]. Therefore, rather than trying to determine, precisely, whether or not the problem is feasible, we seek an approach that is computationally cheap and often finds feasible points for reasonable instances of the problem. The proposed approach involves selecting an initial diagonal power allocation matrix and evaluating each of the integral in (17b). If that power allocation is not feasible, the allocation is iteratively doubled until a feasible allocation is found or the power become unreasonably large. In the latter case a new initial power allocation can be selected and the search for a feasible allocation repeated, or the algorithm reports that no feasible point was found. The question that remains is how to choose the initial power allocation. In our experiments we have found that choosing the initial power allocation to be the power allocation that would be chosen if the channel estimates $\hat{\mathbf{h}}_k^H$ were exact (perfect CSI) and if each SINR_k were set to be equal to its target value, γ_k , typically leads to a feasible starting point for the main algorithm after a small number of the doubling iterations described above. That initial allocation is the solution of the following set of linear equations [42]:

$$\begin{pmatrix} n_1^2 & -m_{12}^2 & \dots & -m_{1K}^2 \\ -m_{21}^2 & n_2^2 & \dots & -m_{2K}^2 \\ \vdots & \vdots & \ddots & \vdots \\ -m_{K1}^2 & -m_{K2}^2 & \dots & n_K^2 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_K \end{pmatrix} = \begin{pmatrix} \sigma_1^2 \\ \sigma_2^2 \\ \vdots \\ \sigma_K^2 \end{pmatrix}, \quad (20)$$

where $m_{ki} = |\hat{\mathbf{h}}_k^H \mathbf{b}_i|$ and $n_k = \frac{1}{\sqrt{\gamma_k}} |\hat{\mathbf{h}}_k^H \mathbf{b}_k|$. A closely related alternative would be to set the initial power allocation so that the “perfect CSI SINRs” are set to a value above γ_k , say $(1 + \Gamma)\gamma_k$, where Γ denotes the (relative) SINR margin; cf. [23], [45]. Among other options, one could also consider using the powers obtained from (13) as the initial point, but computing those powers incurs a significant computational cost.

The proposed robust power allocation algorithm is summarized in Algorithm 1. If the initial power allocation is feasible, and $\Delta_k^{(i)} \rightarrow 0$ (and the integral is computed precisely) the algorithm is guaranteed to converge to a global optimum [14],

³That is, the set of all feasible power allocations

Algorithm 1

-
- 1: Given a feasible diagonal power allocation matrix $\mathbf{P} = \mathbf{P}^{(0)}$, and parameters $\Delta_k^{(0)}$, Δ_k^{\min} , set $i = 0$.
 - 2: **while** $\exists k$ for which (19) with $\Delta_k^{(i)}$ replaced by Δ_k^{\min} is not satisfied **do**
 - 3: $i = i + 1$
 - 4: For all k , choose $\Delta_k^{(i)}$ such that $\Delta_k^{(i)} \leq \Delta_k^{(i-1)}$
 - 5: **for** $k = 1 \rightarrow K$ **do**
 - 6: Using bisection search, find $p_k = p_k^{(i)} \in [0, p_k^{(i-1)}]$ satisfying (19)
 - 7: **end for**
 - 8: **end while**
-

[33]. However, our simple algorithm for finding an initial point does not provide any guarantees. Nevertheless, we will demonstrate in Section VII that by tackling the problem directly, without a conservative approximation, the proposed approach often provides better performance than the existing conservative approaches. Having said that, the repeated requirement to compute an integral of the form in (17b) imposes a significant computational burden. (The SDPs that must be solved in the existing conservative approaches, such as those in (8), also impose a significant computational burden.) To address this issue, in the following sections we will develop customized variants of the algorithm for the case of the zero-forcing directions.

VI. EFFICIENT ALGORITHMS FOR THE ZERO-FORCING CASE

When the nominally zero-forcing directions, $\mathbf{B} = \mathbf{B}_{\text{ZF}} = \hat{\mathbf{H}}^H (\hat{\mathbf{H}} \hat{\mathbf{H}}^H)^{-1}$, are chosen, which implicitly requires $K \leq N_t$, the structure of the integrand in (17b) simplifies, and this simplification facilitates an approximation of the integrand that enables straightforward application of residue theory to obtain an analytic expression for the integral. That analytic expression can be evaluated much more easily than the integral in (17b).

A. Basic formulation for the ZF case

For the case of (nominally) zero-forcing beamforming, $\bar{\mathbf{B}}_k^H \hat{\mathbf{h}}_k = \mathbf{0}$ and $\hat{\mathbf{h}}_k^H \mathbf{b}_k = 1$. These simplifications enable us to rewrite the robust power loading problem in (9) as

$$\min_{\{p_k \geq 0\}} \text{Tr}(\mathbf{B}_{\text{ZF}} \mathbf{P} \mathbf{B}_{\text{ZF}}^H) \quad (21a)$$

$$\text{s.t. } \Pr_{\delta_k} \left(\delta_k^H \mathbf{Q}_k \delta_k + \left(\frac{p_k}{\gamma_k} \right) 2 \text{Re}(\delta_k^H \tilde{\mathbf{r}}_k) + v_k \geq 0 \right) \geq 1 - \epsilon_k, \quad \forall k, \quad (21b)$$

where we have used the form of the chance constraint in (12). While \mathbf{Q}_k takes the same form as in (12), the other parameters simplify to $\tilde{\mathbf{r}}_k = \mathbf{C}_k^{1/2} \mathbf{b}_k$ and $v_k = \frac{p_k}{\gamma_k} - \sigma_k^2$. Note that $\mathbf{r}_k = \frac{p_k}{\gamma_k} \tilde{\mathbf{r}}_k$; this rescaling simplifies the discussion below. Using Lemma 1, each chance constraint in (21b) can be rewritten in the deterministic form

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{\tau_k(i\omega + \beta)}}{i\omega + \beta} \frac{e^{-c_k}}{\det(\mathbf{I} - (i\omega + \beta) \mathbf{Q}_k)} d\omega \geq 1 - \epsilon_k, \quad (22)$$

where $\tau_k = -\sigma_k^2$, $c_k = \sum_{m=1}^M \frac{|\tilde{a}_{km}|^2(i\omega+\beta)\lambda_{mk}}{1+(i\omega+\beta)\lambda_{mk}}$ and $\mathbf{a}_k = -\mathbf{C}_k^{-1/2}\hat{\mathbf{h}}_k$. Here, we let $(-\mathbf{Q}_k) = \mathbf{V}_k\mathbf{\Lambda}_k\mathbf{V}_k^H$ denote the eigen decomposition of $(-\mathbf{Q}_k)$, with λ_{mk} denoting the m^{th} largest eigenvalue, and define $\tilde{\mathbf{a}}_k = \mathbf{V}_k^H\mathbf{a}_k$ with \tilde{a}_{km} being the m^{th} element of $\tilde{\mathbf{a}}_k$.

B. Approximation and residue computation

Since the ℓ^{th} term of the Taylor series expansion of e^{-c_k} has M poles of multiplicity ℓ , it is complicated to apply residue theory directly to the integral in (22). However, if c_k were a constant, then the integrand would take the form of $e^{\tau_k s}G_k(s)$, where $G_k(s)$ is a rational function of $s = \beta + i\omega$. In that case, the application of residue theory is quite straightforward.

One way to approximate the probability in (21b) so that c_k is constant is to approximate the linear term $2\text{Re}(\delta_k^H \tilde{\mathbf{r}}_k)$ by an appropriate constant value, which we will denote by η_k . That is, we approximate the constraint in (21b) by

$$\Pr\left(\delta_k^H \mathbf{Q}_k \delta_k + \left(\frac{p_k}{\gamma_k}\right)\eta_k + v_k \geq 0\right) \geq 1 - \epsilon_k. \quad (23)$$

Since δ_k is a zero-mean Gaussian random variable (RV) with identity covariance matrix, the term $2\text{Re}(\delta_k^H \tilde{\mathbf{r}}_k)$ is a zero-mean Gaussian RV with variance $4\|\tilde{\mathbf{r}}_k\|^2$. That immediately suggests choosing η_k to be a negative multiple of $2\|\tilde{\mathbf{r}}_k\|$. As we explain in Appendix I, the choice of the multiple involves a tradeoff between the accuracy and conservatism of the constraint. Guided by the experiments in [42], we will choose the multiple to be -1.3 in our numerical experiments.

Given the choice of values for each η_k , the robust power loading problem in (21) can be approximated by

$$\min_{\{p_k \geq 0\}} \text{Tr}(\mathbf{B}_{\text{ZF}}\mathbf{P}\mathbf{B}_{\text{ZF}}^H) \quad (24a)$$

$$\text{s.t. } \Pr_{\delta_k}(\delta_k^H \mathbf{Q}_k \delta_k + v'_k \geq 0) \geq 1 - \epsilon_k, \quad (24b)$$

where $v'_k = \frac{p_k}{\gamma_k}\eta_k + v_k = \frac{p_k}{\gamma'_k} - \sigma_k^2$, and $\gamma'_k = \frac{\gamma_k}{1+\eta_k}$. Using Lemma 1, the deterministic equivalent of the chance constraint in (24b) is

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{v'_k(i\omega+\beta)}}{i\omega+\beta} \frac{1}{\det(\mathbf{I} - (i\omega+\beta)\mathbf{Q}_k)} d\omega \geq 1 - \epsilon_k. \quad (25)$$

In terms of the goal of developing an efficient algorithm for robust power loading for the downlink, the key difference between the approximate constraint in (25) and the exact constraint in (22) is that the structure of the numerator of the integrand in (25) enables an application of residue theory to simplify the (exact) computation of that integral. In particular, as shown in Appendix II, when the non-zero eigenvalues of \mathbf{Q}_k are distinct, the integral in (25), and hence the probability on the left hand side of (24b), is

$$\begin{cases} 1 + \sum_{\ell=1}^{K-1} f_{\ell k}(\mathbf{P}) & \text{if } p_k \geq \gamma'_k \sigma_k^2 \\ -f_{rk}(\mathbf{P}) & \text{if } p_k < \gamma'_k \sigma_k^2 \end{cases} \quad (26)$$

where

$$f_{\ell k}(\mathbf{P}) = \begin{cases} 0 & \text{if } \lambda_{\ell k} = 0 \\ -\exp\left(\left(\frac{1}{\gamma'_k}p_k - \sigma_k^2\right)\frac{-1}{\lambda_{\ell k}}\right) \frac{1}{\prod_{j \neq \ell} (1 - \lambda_{jk}/\lambda_{\ell k})} & \text{otherwise} \end{cases} \quad (27)$$

and λ_{mk} is the m^{th} largest eigenvalue of $(-\mathbf{Q}_k)$. As shown in Appendix II, for a large class of channel distributions the non-zero eigenvalues of \mathbf{Q}_k are distinct with high probability.

C. Feasible coordinate descent algorithm for (24)

A straightforward approach to exploiting the above analysis is simply to replace integral calculation that is implicit in Step 5 of Algorithm 1 by (26). Since (26) can be computed with much less effort than (22) this algorithm incurs a much lower computational cost than a direct application of Algorithm 1 to (21). Since the problem in (24) can also be viewed in the framework of standard interference functions [33], if the initial power allocation is feasible (and we solve for the minimal feasible p_k at each step), then the algorithm converges to a globally optimal solution to (24). The discussion in Appendix I guides the choice of η_k so that, with high probability, solving the approximate problem in (24), with the assistance of (26), is equivalent to solving the original problem in (21).

D. Coordinate update algorithm for approximating (24)

Although employing the approximation in (26) of the integral in (22) results in an algorithm that is significantly cheaper than the generic application of Algorithm 1, each step in the bisection search in Step 5 of the approximate algorithm requires, among other things, the eigen decomposition of the current $(-\mathbf{Q}_k)$. In this section we develop an alternate approximate algorithm that enables direct updating of p_k and enables all the powers to be updated using only one eigen decomposition of each $(-\mathbf{Q}_k)$. Unlike the algorithms in the previous sections, the iterates of the algorithm in this section are not necessarily feasible, but the power allocation is cyclically updated in such a way that it often converges to a good solution. For that reason, we will refer to the algorithm in this section as a coordinate update algorithm, as distinct from the coordinate descent algorithm in the previous section.

To develop an approximate cyclic coordinate update algorithm for the problem in (24), we begin by observing that $f_{\ell k}(\mathbf{P})$ in (26) depends on p_k both explicitly, and implicitly through the eigenvalues of $(-\mathbf{Q}_k)$. To avoid the complexity that this implicit dependence incurs, in the i^{th} cycle of updates, we will employ the following approximation of $\mathbf{Q}_k^{(i)}$,

$$\hat{\mathbf{Q}}_k^{(i)} = \mathbf{C}_k^{1/2} \left(\frac{p_k^{(i-1)}}{\gamma_k} \mathbf{b}_k \mathbf{b}_k^H - \bar{\mathbf{B}}_k \mathbf{P}^{(i-1)} \bar{\mathbf{B}}_k^H \right) \mathbf{C}_k^{1/2}, \quad (28)$$

where $\mathbf{P}^{(i-1)} = \text{Diag}(p_1^{(i-1)}, p_2^{(i-1)}, \dots, p_K^{(i-1)})$ is the power allocation at the end of the previous cycle.

With that approximation of $\mathbf{Q}_k^{(i)}$ in place, at the k^{th} step of the i^{th} cycle we are looking for a power $p_k^{(i)}$ that lies close to the boundary of the feasible set. That is we are looking for the smallest non-negative $p_k^{(i)}$ such that

$$\begin{cases} 1 + \sum_{\ell=1}^{K-1} \hat{f}_{\ell k}(p_k^{(i)}) \geq 1 - \epsilon_k & \text{if } p_k^{(i)} \geq \gamma'_k \sigma_k^2 \\ -\hat{f}_{rk}(p_k^{(i)}) \geq 1 - \epsilon_k & \text{if } p_k^{(i)} < \gamma'_k \sigma_k^2 \end{cases} \quad (29)$$

where

$$\hat{f}_{\ell k}(p_k^{(i)}) = \begin{cases} 0 & \text{if } \hat{\lambda}_{\ell k}^{(i)} = 0 \\ -\exp\left(\left(\frac{1}{\gamma_k'} p_k^{(i)} - \sigma_k^2\right) \frac{-1}{\hat{\lambda}_{\ell k}^{(i)}}\right) \frac{1}{\prod_{j \neq \ell} (1 - \hat{\lambda}_{jk}^{(i)} / \hat{\lambda}_{\ell k}^{(i)})} & \text{otherwise} \end{cases} \quad (30)$$

and $\hat{\lambda}_{mk}^{(i)}$ is the m^{th} largest eigenvalue of $(-\hat{\mathbf{Q}}_k^{(i)})$. Here we have assumed that the non-zero eigenvalues are distinct.

Since we are looking for the smallest $p_k^{(i)}$, we first consider the second case in (29) where $p_k^{(i)} < \gamma_k' \sigma_k^2$. In this case we look for the smallest $p_k^{(i)}$ that satisfies $\hat{f}_{rk}(p_k^{(i)}) \geq 1 - \epsilon_k$, namely

$$\tilde{p} = \gamma_k' \sigma_k^2 - \gamma_k' \hat{\lambda}_{rk}^{(i)} \ln\left((1 - \epsilon_k) \prod_{j \neq r} (1 - \hat{\lambda}_{jk}^{(i)} / \hat{\lambda}_{rk}^{(i)})\right). \quad (31)$$

If $\tilde{p} \in (0, \gamma_k' \sigma_k^2)$, it is an admissible solution for (29).

If \tilde{p} is not admissible, the desired solution is the smallest non-negative root of $\sum_{\ell=1}^{r-1} \hat{f}_{\ell k}(p_k^{(i)}) + \epsilon_k = 0$ that is not smaller than $\gamma_k' \sigma_k^2$; cf. (29). Since $\hat{f}_{\ell k}(p_k^{(i)})$ is smooth, any one of a number of standard root finding algorithms could be considered. Instead of doing that, we will employ a conservative approximation of the constraint $1 + \sum_{\ell=1}^{r-1} \hat{f}_{\ell k}(p_k^{(i)}) \geq 1 - \epsilon_k$ and show that the resulting problem has a closed-form solution. As ℓ increases, the argument of the exponential in (26) becomes more negative, and hence the magnitude of $\hat{f}_{\ell k}(p_k^{(i)})$ decreases. Furthermore, for odd ℓ , $\hat{f}_{\ell k}(p_k^{(i)}) < 0$, whereas for even ℓ , $\hat{f}_{\ell k}(p_k^{(i)}) > 0$. As a result we have that $\sum_{\ell=2}^{r-1} \hat{f}_{\ell k}(p_k^{(i)}) \geq 0$. (Typically, this term will also be small in comparison to $|\hat{f}_{1k}(p_k^{(i)})|$). Therefore, if $p_k^{(i)}$ is chosen such that $1 + \hat{f}_{1k}(p_k^{(i)}) \geq 1 - \epsilon_k$, then the outage constraint is guaranteed to hold. More explicitly, if we let

$$\check{p} = \gamma_k \sigma_k^2 - \gamma_k \hat{\lambda}_{1k}^{(i)} \ln\left(\epsilon_k \prod_{j \neq 1} (1 - \hat{\lambda}_{jk}^{(i)} / \hat{\lambda}_{1k}^{(i)})\right), \quad (32)$$

then if \check{p} is not admissible we choose $p_k^{(i)} = \max\{\check{p}, \gamma_k' \sigma_k^2\}$.

Having established the iterations above, we need to select the initial powers. Using the insight developed in [46] for a slightly different system, we will chose $p_k^{(0)}$ as if we have equal power allocation with the k^{th} user's parameters. That is, to set $p_k^{(0)}$ we let $\mathbf{P}^{(0)}$ take the form $\check{p}\mathbf{I}$, and determine the value of \check{p} that yields equality in the k^{th} user's approximation of the outage constraint; cf. (24b). This value can be computed in closed form:

$$p_k^{(0)} = \frac{\sigma_k^2}{1/\gamma_k' + \tilde{\lambda}_{1k} \ln(\epsilon_k \prod_{j \neq 1} (1 - \tilde{\lambda}_{jk} / \tilde{\lambda}_{1k}))}, \quad (33)$$

where $\tilde{\lambda}_{mk}$ is the m^{th} largest eigenvalue of the corresponding $-\hat{\mathbf{Q}}_k = -\mathbf{C}_k^{1/2} \left(\frac{1}{\gamma_k} \mathbf{b}_k \mathbf{b}_k^H - \tilde{\mathbf{B}}_k \tilde{\mathbf{B}}_k^H \right) \mathbf{C}_k^{1/2}$. Unlike the coordinate descent algorithm in previous chapter and that in Section VI-C, this initial power allocation is not necessarily feasible, but as the coordinates are updated, the power allocation tends to move toward the feasible set. The cyclic updates are terminated once a feasible point is found or if no feasible point is found in a reasonable time. (Feasibility is evaluated

using the expression in (26) and comparing with $1 - \epsilon_k$.) Our numerical experience suggests that the starting point in (33) is particularly effective in that the level of conservatism in the first feasible point tends to be low. In systems where that is not the case, one can use this feasible point to initialize the algorithm in Section VI-C. The algorithm developed in this section is summarized in Algorithm 2.

Algorithm 2

- 1: Select i_{\max} , and set each $p_k^{(0)}$ according to (33). Set $i = 0$.
 - 2: **while** current power allocation is infeasible and $i \leq i_{\max}$ **do**
 - 3: $i = i + 1$
 - 4: For all k , construct $\hat{\mathbf{Q}}_k^{(i)}$ according to (28) and compute its eigen decomposition.
 - 5: **for** $k = 1 \rightarrow K$ **do**
 - 6: Compute \tilde{p} in (31).
 - 7: **if** $\tilde{p} \in (0, \gamma_k' \sigma_k^2)$ **then** set $p_k^{(i)} = \tilde{p}$
 - 8: **else** compute \check{p} in (32) and set $p_k^{(i)} = \max\{\check{p}, \gamma_k' \sigma_k^2\}$.
 - 9: **end if**
 - 10: **end for**
 - 11: Evaluate feasibility of current power allocation using (24b) and (26).
 - 12: **end while**
-

VII. PERFORMANCE EVALUATION

In this section, we demonstrate the performance of the proposed algorithms. For the general algorithm developed in Section V (see Algorithm 1), we will use regularized channel inversion (RCI) beamforming, $\mathbf{B}_{\text{RCI}} = \hat{\mathbf{H}}^H (\hat{\mathbf{H}} \hat{\mathbf{H}}^H + \alpha \mathbf{I}_K)^{-1}$, and zero-forcing beamforming, $\mathbf{B}_{\text{ZF}} = \hat{\mathbf{H}}^H (\hat{\mathbf{H}} \hat{\mathbf{H}}^H)^{-1}$, and the (PCSI) beamformers that are obtained by treating the estimated CSI as if it were perfect and solving (5), as examples of fixed-direction beamformers. For the RCI case, we specify the regularization parameter, α , according to the results of [41] for the full CSI case, namely $\alpha = K\sigma^2$ where σ^2 is the noise variance at each receiver (which will be assumed to be the same). For the algorithms proposed in Section VI, which seek good solutions to the approximated problem in (24) for the nominally zero-forcing case (see Section VI-C and Algorithm 2), based on analysis in [42] the value of the parameter η_k is chosen to be $-1.3(2\|\tilde{\mathbf{r}}_k\|)$.

We will compare our algorithms to several existing methods, including the chance-constrained robust precoding method of [28], which is based on rank-one relaxation and convex restriction (RAR) and was formulated in (8). This method is a full precoder design method, and hence the beamformers and the power allocation are designed jointly. In terms of power allocation methods, we will compare against the adaptation of the RAR approach to robust power loading of that was formulated in (13), and the MSE-based fixed-point iteration method in [27]. For convenience, a summary of methods that will be considered, and the labels that will describe them, is provided in Table I.

TABLE I
DESCRIPTION OF METHODS

Method	Description
PCSI-General	Robust power loading (RPL) based on the original formulation for the “perfect CSI” directions. Solved using Algorithm 1.
RCI-General	RPL based on the original formulation for the RCI directions with $\alpha = K\sigma^2$. Solved using Algorithm 1.
ZF-General	RPL based on the original formulation for the ZF directions. Solved using Algorithm 1.
ZF-CoordDescent	RPL for the ZF directions based on the algorithm in Section VI-C, which seeks good solutions to the approximated problem in (24) with $\eta_k = -1.3(2\ \tilde{\mathbf{r}}_k\)$.
ZF-CoordUpdate	RPL for the ZF directions based on Algorithm 2, which seeks good solutions to the approximated problem in (24) with $\eta_k = -1.3(2\ \tilde{\mathbf{r}}_k\)$.
RAR	Robust precoding based on the rank relaxation and convex restriction (RAR) method of [28]. The resulting SDP was formulated in (8).
ZF-SDP	RPL for the ZF directions based on the adaptation of the RAR method of [28] to the power loading problem that was formulated in (13).
Fixed-point	RPL for the ZF directions based on the fixed-point approach in [27].

In the following simulations, we consider an environment with $N_t = 3$ transmit antennas, $K = 3$ users, i.i.d. Rayleigh fading channels, and the receivers’ noise sources are modeled as zero-mean, additive, white, and Gaussian with variance $\sigma_k^2 = \sigma^2 = 0.01$. In our simulations the errors in the BS’s estimates of the channels are generated by explicitly performing uplink training with orthogonal training sequences and linear MMSE channel estimation [36].⁴ Even though the linear MMSE estimator (of a given realization of the channel) is only asymptotically unbiased (as the training SNR increases), for the purposes of robust power loading the uncertainties in the channel estimates will be modeled by Gaussian random vectors with zero mean. The error covariance for the linear MMSE estimator in this setting is $\mathbf{C}_k = \mathbf{C} = \sigma_e^2 \mathbf{I}$, where $\sigma_e^2 = \frac{\sigma_{\text{BS}}^2}{\sigma_{\text{BS}}^2 + L_{\text{UT}} P_{\text{UT}}}$, with σ_{BS}^2 being the noise variance at the BS’s receiver for the uplink, L_{UT} being the length of the training sequence, and P_{UT} being the power of the training sequence, and this covariance will be used in the model of the error. In terms of performance specifications, the probability of outage is set to be $\epsilon_k = \epsilon = 0.05$ for all users and a universal SINR target is defined; i.e., $\gamma_k = \gamma$.

A. Performance Comparisons Against SINR Requirements

In this section, we randomly generated 10,000 realizations of the set of i.i.d. Rayleigh fading channels $\{\mathbf{h}_k^H\}_{k=1}^K$. In the uplink training phase, we assumed that there were no peak power constraints on the uplink and we set $L_{\text{UT}} = 1$ and $P_{\text{UT}} = 4.99$ so that with the noise variance at the base station

⁴The TDD “ping-pong” time is assumed to be short enough and the RF calibration good enough (cf. [38]) for the base-band equivalent channels to be reciprocal.

being $\sigma_{\text{BS}}^2 = 0.01$, the variance of the channel estimate is $\sigma_e^2 = 0.002$. (The choice of $L_{\text{UT}} = 1$ is consistent with the results of [47].) We examined the performance of each design method as the SINR requirement of the users, γ , increases from 0 dB to 10 dB. For each set of channel estimates and for each value of γ , we determined whether each method yields a successful design, in the sense that it generates a precoder that satisfies the outage constraints in the original formulation in (7). (For the power loading algorithms, the formulation in (9) is equivalent.) Satisfaction of the constraints was determined by evaluating the outage probability using Lemma 1; cf. (17b). In other words, to be deemed “successful” for a given set of channel estimates, a design method must produce a solution that is not only feasible for the optimization problem on which the design is based, but is also feasible for the original robust precoding problem in (7).⁵ In Fig. 2, for each design method we plot the percentage of channel realizations for which a successful design was obtained.⁶ In Fig. 3, we plot, against γ , the average transmission power over the 7,107 channel set realizations for which all methods provided a successful design at all the considered SINR targets.

From Fig. 2, it can be seen that by tackling the power loading problem directly (or closely), the proposed power loading methods are able to satisfy the QoS constraints more often than the existing power loading method that is based on an optimal solution to a tractable conservative approximation of the problem (ZF-SDP); cf. (13). This is because the approximation made in that method can be quite conservative.⁷ That conservatism can also be seen in Fig. 3: For the channel realizations for which all methods produce a successful design, the ZF-SDP method uses a larger transmission power than the proposed methods. Figs 2 and 3 also show that the proposed methods satisfy the QoS constraints more often than the fixed-point method in [27], and that they expend less power in doing so.

Among the “General” methods, it is interesting to observe the role that the choice of beamforming directions has on the performance. When the SINR targets are low, the “perfect CSI” directions provide the best performance of all the methods we have considered, but when the SINR targets are higher, the nominally ZF directions provide the best performance. Furthermore, at higher SINR targets the combinations of the ZF

⁵More specifically, when the proposed General algorithm produces a feasible solution to (17), that solution is guaranteed to be feasible for (7). Similarly, when the ZF-SDP method produces a feasible solution to (13), and when the Fixed-point algorithm produces a solution, those solutions are also guaranteed to be feasible for (7). For the RAR precoding method, a feasible solution to (8) in which all the matrices \mathbf{U}_k have rank one generates a feasible solution to (7). (If not all of those matrices have rank one, then one can use a variety of techniques to try to generate a feasible solution to (7) from the solution to (8).) For the ZF-CoordDescent and ZF-CoordUpdate methods, the use of the approximation in (23) to obtain significant reductions in the computational cost means that when these algorithms produce feasible solutions to (24), one has to test whether that solution is feasible for (7).

⁶Since the training SNR on the uplink is reasonably high, the corresponding results for least squares channel estimation, which is the minimum variance unbiased estimator in this setting, are indistinguishable from the results in Fig. 2 at the scale of that figure.

⁷The approximation made in the ZF-SDP method involves approximating the outage constraint by the constraint that the design guarantee zero outage for all uncertainties up to a specified size, where that size is chosen so that the probability of larger uncertainties is less than the specified outage probability.

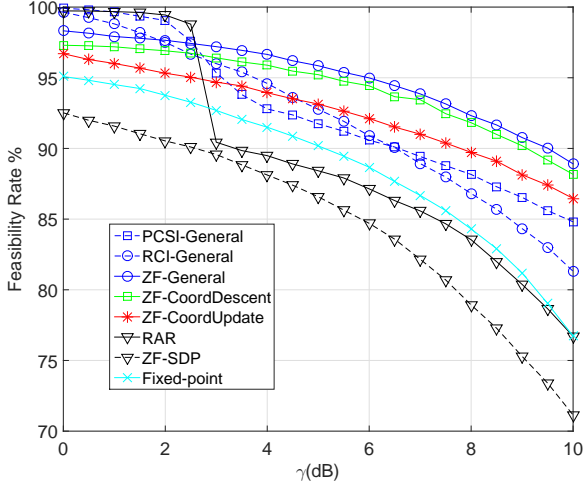


Fig. 2. Percentage of successful designs for different methods in the environment where $N_t = K = 3$, $\mathbf{C} = 0.002\mathbf{I}$, $\epsilon = 0.05$, and $\sigma^2 = 0.01$.

directions and the computationally more efficient approximate power loading algorithms provide better performance than the combinations of the PCSI or RCI directions and the General power loading algorithm.

What is perhaps more interesting is that for higher SINR targets, the proposed power loading methods provide better performance than the RAR robust precoding method [28], which is formulated in (8), despite the fact that the RAR method has many more degrees of design freedom.⁸ Once again, this is due to the fact that the techniques used in the RAR method to convert the chance constraints into deterministic constraints can be quite conservative.⁹ (An explicit example of the reduced conservatism of the proposed methods is available in [42].)

We have performed analogous experiments systems with more antennas and users (4, 5 and 6), and the general structure of Figs 2 and 3 and the observations made in the previous two paragraphs apply in those cases, too.

B. Performance Comparisons Against Uncertainty Size

In the experiments in this section, we randomly generated 1,000 realizations of the set of i.i.d. Rayleigh fading channels $\{\mathbf{h}_k^H\}_{k=1}^K$, and similar to the previous simulation, we obtained the CSI at the BS through uplink training. In the uplink we choose $L_{UT} = 1$, and chose P_{UT} to vary the variance of the channel uncertainty, σ_e^2 . In this experiment we set the SINR target of the users to 3 dB. In Fig. 4 we plot the percentage of channel set realizations for which each algorithm generates a successful design, as P_{UT} is decreased (and hence

⁸The RAR method designs the beamformers and power allocation jointly, whereas in the proposed methods, the ZF-SDP method, and the Fixed-point method, the beamformers are fixed.

⁹Like the ZF-SDP method, the RAR method is based on a “zero-outage region” approximation of the outage constraint. For lower SINR targets the extra degrees of design freedom in the RAR method overcome the conservatism in this approximation, but for higher SINR targets the extra degrees of design freedom only provide small performance gains over the ZF-SDP method.

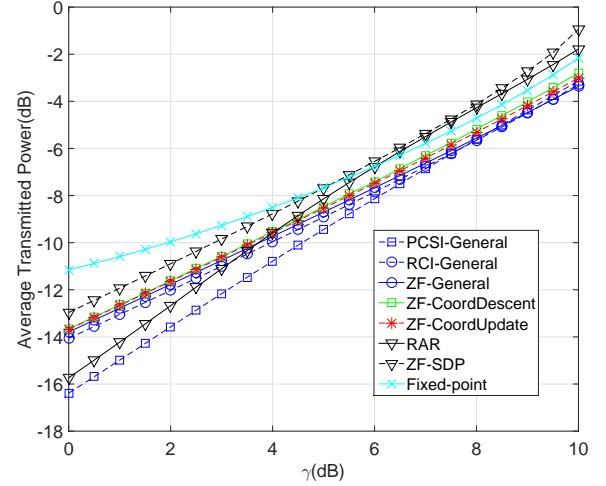


Fig. 3. Power transmission performance for different methods in the environment where $N_t = K = 3$, $\mathbf{C} = 0.002\mathbf{I}$, $\epsilon = 0.05$, and $\sigma^2 = 0.01$; i.e., the receiver noise power is -20dB .

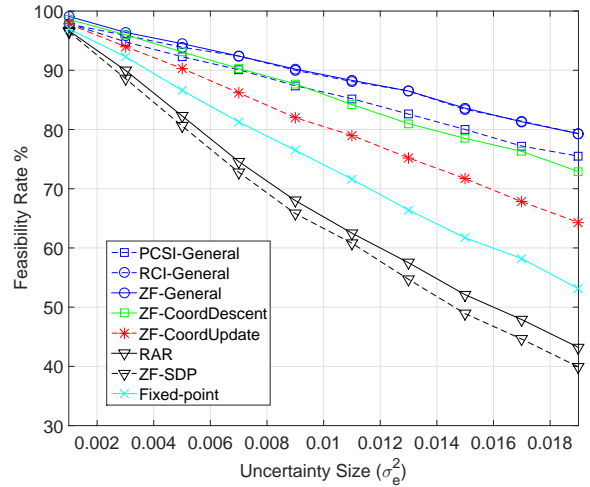


Fig. 4. Percentage of successful designs for different methods in the environment where $N_t = K = 3$, $\gamma = 3\text{ dB}$, $\epsilon = 0.05$, and $\sigma^2 = 0.01$.

σ_e^2 increases). In Fig. 5 we plot the average power that each algorithm requires over the 400 channel set realizations for which all methods provide a successful design for all the considered values of σ_e^2 . Figs 4 and 5 show that as the variance of the channel uncertainty increases, the performance of the proposed methods degrades much more slowly than that of the existing methods.

C. Computational Cost Comparisons

As illustrated in Figs 2–5, among the proposed algorithms, the “General” algorithms provide the best performance. This is to be expected because they tackle the original problem without any approximation. However, these algorithms are also the most computationally expensive of the proposed algorithms. In the i^{th} cycle they require the evaluation of $\sum_k N_{\text{bisect},k}^{(i)}$ integrals of the form in (19), where $N_{\text{bisect},k}^{(i)}$ is

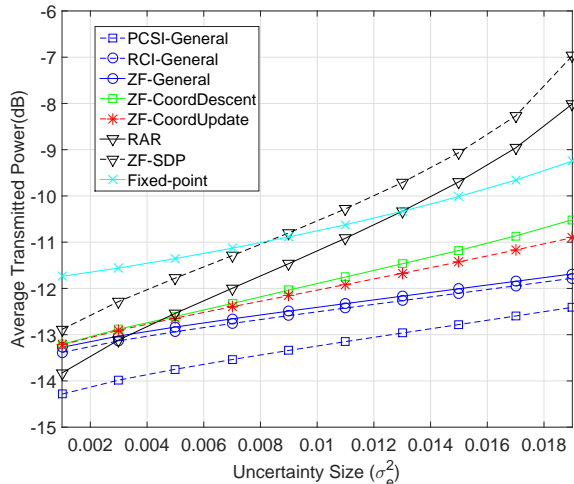


Fig. 5. Power transmission performance for different methods in the environment where $N_t = K = 3$, $\gamma = 3$ dB, $\epsilon = 0.05$, and $\sigma^2 = 0.01$; i.e., the receiver noise power is -20 dB.

the number of bisection steps needed to update p_k in the i^{th} cycle. Each of those integrals involves the computation of an eigen decomposition of a Hermitian symmetric matrix of size $N_t \times N_t$, which requires $O(N_t^3)$ operations, and the evaluation of a scalar indefinite integral. To evaluate the total computational cost of this algorithm we also need to analyze $N_{\text{bisection},k}^{(i)}$ and the number of coordinate descent cycles. This is a substantially more difficult task, but we can say that in our numerical experiments for SINR targets of 5 dB, the median of the total number of integrals to be computed, $\sum_i \sum_k N_{\text{bisection},k}^{(i)}$, was 21 for the RCI beamformers and 22 for the ZF beamformers. Furthermore, in these MATLAB-based numerical experiments, the run-times of our General algorithms were of similar magnitude to those of the RAR and ZF-SDP methods, although our implementations of the RAR and ZF-SDP methods used compiled code to solve the SDPs. As a more analytical comparison, we observe that each interior point iteration in the SDP solver requires $O(N_t^6)$ operations.

The ZF-CoordDescent algorithm is significantly faster than the General methods, with the key computational task in each cycle being the computation of $\sum_k N_{\text{bisection},k}^{(i)}$ eigen decompositions, each of which requires $O(N_t^3)$ operations. In our numerical experiments for SINR targets of 5 dB, the median of the total number of bisection steps, $\sum_i \sum_k N_{\text{bisection},k}^{(i)}$, was 59. Although that number is larger than in the General case, the cost per bisection step is smaller because there is no integral to calculate. The cost-per-cycle of the ZF-CoordUpdate algorithm is significantly smaller than that of the ZF-CoordDescent algorithm, with the key task in each cycle being the computation of only K eigen decompositions, each of which requires $O(N_t^3)$ operations. In our numerical experiments for SINR targets of 5 dB, the median number of cycles of the ZF-CoordUpdate algorithm was only 2. In those MATLAB-based experiments, the run-times of the ZF-CoordDescent and ZF-CoordUpdate methods were about an order of magnitude lower than that of the “General”, RAR

and ZF-SDP methods.

Of the methods considered, the fixed-point method was the fastest in our experiments. Even though its computational cost per cycle is also $O(N_t^3)$, and the median number of cycles for SINR targets of 5 dB was 19, in practice it was significantly faster than the proposed ZF-CoordUpdate algorithm. That said, proposed algorithm does provide somewhat better performance, and expends less power in doing so.

VIII. CONCLUSION

In this paper we have developed algorithms for robust power loading in the MISO downlink beamforming scenario in the presence of Gaussian uncertainties in the base station’s estimates of the channels. These algorithms are well-suited to systems that operate in a TDD manner. In contrast to many of the existing approaches, the proposed algorithms are based on a precise deterministic characterization of the probability of outage that is induced by the uncertainty, rather than conservative deterministic characterizations. The precise characterization was incorporated into a cyclic coordinate descent algorithm that can be viewed within the framework of standard interference functions and hence is guaranteed to converge to a globally optimal solution. Insight into the deterministic characterization then led to the development of approximations that yield much faster algorithms that retain much of the performance of the original algorithm. The performance of the proposed algorithms was examined in a number of simulation studies, and they were shown to have highly desirable performance characteristics, especially in cases of higher SINR targets and larger uncertainties. In particular, the proposed algorithms provided significantly better performance than those based on conservative approximations of the impact of the uncertainty.

The proposed robust power loading techniques have been developed and evaluated in the context of a single-cell downlink with a modest number of transmitter antennas, but they can be extended in a straightforward way to the multi-cell downlink scenario with centralized precoder design. In principle, they can also be extended to the “Massive MIMO” downlink scenario, but for reasons of computational cost such extensions would likely focus on the class of hybrid analog-digital beamforming strategies for Massive MIMO; e.g., [48].

APPENDIX I CHOICE OF η_k

One side of the tradeoff in the choice of η_k arises from the observation that the probability on the left hand side of (23) is a increasing function of η_k . The other side arises from the observation that $\rho_k(\eta_k) = \Pr(2 \text{Re}(\delta_k^H \tilde{\mathbf{r}}_k) \geq \eta_k)$ is a decreasing function of η_k . It is tempting to consider choosing η_k so that $\rho_k(\eta_k)$ is large, as this increases the probability that designing a power loading that satisfies the approximated QoS constraint in (23) produces a power loading that satisfies the original QoS constraint in (21b). However, doing so is inherently conservative. In particular, the approximated problem maybe infeasible when the original problem is, in fact, feasible.

Rather than fixing η_k to a particular value, an alternative approach is to change it iteratively based on an evaluation of whether or not the current power loading satisfies the original QoS constraint [42]. Doing so provides a slight performance improvement in terms of the feasibility rate, and a significant reduction in the power allocated when a feasible solution is found [42], but incurs the additional cost of computing the integrals in (22) in each cycle.

APPENDIX II DERIVATION OF (26)

To prove the statement in (26) and (27), we observe that

$$\begin{aligned} & \Pr(\|\delta_k\|_{(-\mathbf{Q}_k)}^2 \leq p_k/\gamma'_k - \sigma_k^2) \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{(p_k/\gamma'_k - \sigma_k^2)(i\omega + \beta)}}{i\omega + \beta} \frac{1}{\det(\mathbf{I} + (i\omega + \beta)(-\mathbf{Q}_k))} d\omega \\ &= \frac{1}{2\pi i} \int_{-i\infty + \beta}^{i\infty + \beta} \frac{e^{(p_k/\gamma'_k - \sigma_k^2)s}}{s} \frac{1}{\prod_j (1 + s\lambda_{jk})} ds. \end{aligned} \quad (34)$$

where λ_{jk} is the j^{th} eigenvalue of $-\mathbf{Q}_k$, and, as defined after (12), $(-\mathbf{Q}_k) = \mathbf{C}_k^{1/2} (-\frac{p_k}{\gamma_k} \mathbf{b}_k \mathbf{b}_k^H + \bar{\mathbf{B}}_k \mathbf{P} \bar{\mathbf{B}}_k^H) \mathbf{C}_k^{1/2}$, with \mathbf{b}_k and $\bar{\mathbf{B}}_k$ containing the appropriate columns of the matrix of nominally zero-forcing directions, \mathbf{B}_{ZF} . (Recall that $K \leq N_t$ in this case.) When some power is allocated to each user (i.e., when each $p_k > 0$), $(-\mathbf{Q}_k)$ is the sum of a positive semi-definite matrix of generic rank $(K-1)$ and a rank one negative definite matrix. The structure of these matrices and the nature of the zero-forcing directions means that for a large class of channel distributions, with high probability $(-\mathbf{Q}_k)$ has $K-1$ distinct positive eigenvalues and one negative eigenvalue.

We will denote the integrand in (34) by $F(s)$, and in Fig. 6, we have illustrated the generic shape of the location of the poles of $F(s)$. The integral in (34) that we are trying to evaluate is along the vertical line with real value β . According to Lemma 1, β can be any positive constant value that satisfies $1 + \beta\lambda_{mk} > 0$ for all m . For positive eigenvalues this constraint always holds, but for the negative eigenvalue, λ_{rk} , we must have $\beta < \frac{-1}{\lambda_{rk}}$. This condition and the fact that $\beta > 0$ require that β be chosen in the shaded region of Fig. 6. The next step in applying residue theory is to pick the appropriate closed contour such that the integral in (34) is equal to integral along that contour. For this purpose, a path should be added to complete the closed contour in such a way that the integral on the added path is zero. To find the appropriate path to be added, the problem is divided into two cases. In the case that $p_k/\gamma'_k - \sigma_k^2 \geq 0$, the suitable contour, C_1 , is depicted on the left of Fig. 6, since the integrand is zero for the semicircular portion of the path. Therefore, when the positive eigenvalues of $(-\mathbf{Q}_k)$ are distinct we can use residue theory to simplify the integral in (34) to

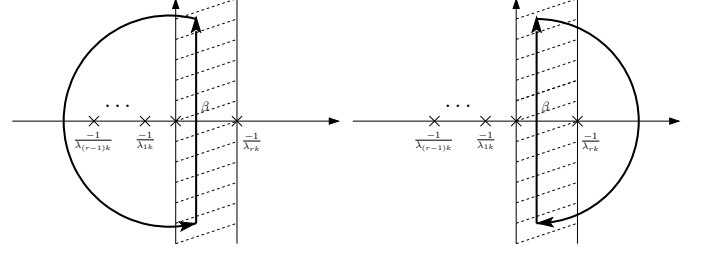


Fig. 6. Appropriate contour for the cases that $p_k/\gamma'_k - \sigma_k^2 \geq 0$ (left) and $p_k/\gamma'_k - \sigma_k^2 < 0$ (right).

$$\begin{aligned} \frac{1}{2\pi i} \oint_{C_1} F(s) ds &= \text{Res}_{s=0} F(s) + \sum_{\ell=1}^{K-1} \text{Res}_{s=-\lambda_{\ell k}^{-1}} F(s) \\ &= 1 + \sum_{\ell=1}^{K-1} \lim_{s \rightarrow -\lambda_{\ell k}^{-1}} (s + \lambda_{\ell k}^{-1}) \frac{e^{(p_k/\gamma'_k - \sigma_k^2)s}}{s} \frac{1}{\prod_j (1 + s\lambda_{jk})} \\ &= 1 + \sum_{\ell=1}^{K-1} f_{\ell k}(\mathbf{P}), \end{aligned} \quad (35)$$

where $f_{\ell k}(\cdot)$ was defined in (27).

On the other hand, for the case that $p_k/\gamma'_k - \sigma_k^2 < 0$, the suitable contour, C_2 , is depicted on the right of Fig. 6, since the integrand is zero for semicircular portion of that path. So in this case the residue theory should be applied just one pole. Note that this time the contour is clockwise, so the integral along the contour is equal to the negative of the residue. In this setting the integral in (34) can be simplified as follows

$$\begin{aligned} \frac{1}{2\pi i} \oint_{C_2} F(s) ds &= -\text{Res}_{s=-\lambda_{rk}^{-1}} F(s) \\ &= -\lim_{s \rightarrow -\lambda_{rk}^{-1}} (s + \lambda_{rk}^{-1}) \frac{e^{(p_k/\gamma'_k - \sigma_k^2)s}}{s} \frac{1}{\prod_j (1 + s\lambda_{jk})} \\ &= -f_{rk}(\mathbf{P}). \end{aligned} \quad (36)$$

REFERENCES

- [1] J. H. Winters, J. Salz, and R. D. Gitlin, "The impact of antenna diversity on the capacity of wireless communication systems," *IEEE Trans. Commun.*, vol. 42, no. 2/3/4, pp. 1740–1751, Feb./Mar./Apr. 1994.
- [2] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936–3964, Sep. 2006.
- [3] D. Gesbert, M. Kountouris, R. W. Heath, C.-B. Chae, and T. Salzer, "Shifting the MIMO paradigm," *IEEE Signal Process. Mag.*, vol. 24, no. 5, pp. 36–46, 2007.
- [4] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1691–1706, 2003.
- [5] C. Windpassinger, R. F. Fischer, T. Vencel, and J. B. Huber, "Precoding in multiantenna and multiuser communications," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1305–1316, 2004.
- [6] C.-H. Fung, W. Yu, and T. J. Lim, "Precoding for the multiantenna downlink: Multiuser SNR gap and optimal user ordering," *IEEE Trans. Commun.*, vol. 55, no. 1, pp. 188–197, 2007.
- [7] J. Liu and W. A. Krzymien, "A novel nonlinear joint transmitter-receiver processing algorithm for the downlink of multiuser MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 57, no. 4, pp. 2189–2204, 2008.
- [8] B. M. Hochwald, C. B. Peel, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication—Part II: Perturbation," *IEEE Trans. Wireless Commun.*, vol. 53, no. 3, pp. 537–544, 2005.

- [9] F. Rashid-Farrokhi, L. Tassiulas, and K. J. R. Liu, "Joint optimal power control and beamforming in wireless networks using antenna arrays," *IEEE Trans. Commun.*, vol. 46, no. 10, pp. 1313–1324, Oct. 1998.
- [10] M. Bengtsson and B. Ottersten, "Optimal and suboptimal transmit beamforming," in *Handbook of Antennas in Wireless Communications*, L. C. Godara, Ed. CRC Press, 2001, ch. 18.
- [11] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 461–471, 2004.
- [12] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Trans. Veh. Tech.*, vol. 53, no. 1, pp. 18–28, Jan. 2004.
- [13] A. Wiesel, Y. Eldar, and S. Shamai, "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 161–176, Jan. 2006.
- [14] M. Schubert and H. Boche, "A generic approach to QoS-based transceiver optimization," *IEEE Trans. Commun.*, vol. 55, no. 8, pp. 1557–1566, Aug. 2007.
- [15] R. Hunger and M. Joham, "A complete description of the QoS feasibility region in the vector broadcast channel," *IEEE Trans. Signal Process.*, vol. 58, no. 7, pp. 3870–3878, Jul. 2010.
- [16] Y. Huang and D. P. Palomar, "Rank-constrained separable semidefinite programming with applications to optimal beamforming," *IEEE Trans. Signal Process.*, vol. 58, no. 2, pp. 664–678, Feb. 2010.
- [17] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser MIMO achievable rates with downlink training and channel state feedback," *IEEE Trans. Info. Theory*, vol. 56, no. 6, pp. 2845–2866, Jun. 2010.
- [18] J. Xu, J. G. Andrews, and S. A. Jafar, "MISO broadcast channels with delayed finite-rate feedback: Predict or observe?" *IEEE Trans. Wireless Commun.*, vol. 11, no. 4, pp. 1456–1467, Apr. 2012.
- [19] M. Botros Shenouda and T. N. Davidson, "Convex conic formulations of robust downlink precoder design with quality of service constraints," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 4, pp. 714–724, Dec. 2007.
- [20] G. Zheng, K.-K. Wong, and T.-S. Ng, "Robust linear MIMO in the downlink: A worst-case optimization with ellipsoidal uncertainty regions," *EURASIP J. Adv. Signal Process.*, vol. 2008, no. 609018, Jul. 2008.
- [21] N. Vucic and H. Boche, "Robust QoS-constrained optimization of downlink multiuser MISO systems," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 714–725, Feb. 2009.
- [22] M. Botros Shenouda and T. N. Davidson, "Non-linear and linear broadcasting with QoS requirements: Tractable approaches for bounded channel uncertainties," *IEEE Trans. Signal Process.*, vol. 57, no. 5, pp. 1936–1947, May 2009.
- [23] S. Kandukuri and S. Boyd, "Optimal power control in interference-limited fading wireless channels with outage-probability specifications," *IEEE Trans. Wireless Commun.*, vol. 1, no. 1, pp. 46–55, Jan. 2002.
- [24] B. Chalise, S. Shahbazpanahi, A. Czylik, and A. B. Gershman, "Robust downlink beamforming based on outage probability specifications," *IEEE Trans. Wireless Commun.*, vol. 6, no. 10, pp. 3498–3503, 2007.
- [25] M. Payaró, A. Pascual-Iserte, and M. Á. Lagunas, "Robust power allocation designs for multiuser and multi-antenna downlink communication systems through convex optimization," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 7, pp. 1390–1401, Sep. 2007.
- [26] M. Botros Shenouda and T. N. Davidson, "Probabilistically-constrained approaches to the design of the multiple antenna downlink," in *Conf. Rec. 42nd Ann. Asilomar Conf. Signals, Systems, Computers*, Pacific Grove, CA, Oct. 2008, pp. 1120–1124.
- [27] N. Vucic and H. Boche, "A tractable method for chance-constrained power control in downlink multiuser MISO systems with channel uncertainty," *IEEE Sig. Proc. Lett.*, vol. 16, no. 5, pp. 346–349, 2009.
- [28] K.-Y. Wang, A. M.-C. So, T.-H. Chang, W.-K. Ma, and C.-Y. Chi, "Outage constrained robust transmit optimization for multiuser MISO downlinks: Tractable approximations by conic optimization," *IEEE Trans. Signal Process.*, vol. 62, no. 21, pp. 5690–5705, Nov. 2014.
- [29] C. Kohninakis, C. Fragouli, A. H. Sayed, and R. D. Wesel, "Multi-input multi-output fading channel tracking and equalization using Kalman estimation," *IEEE Trans. Signal Process.*, vol. 50, no. 5, pp. 1065–1076, May 2002.
- [30] Z. Liu, X. Ma, and G. B. Giannakis, "Space-time coding and Kalman filtering for time-selective fading channels," *IEEE Trans. Commun.*, vol. 50, no. 2, pp. 183–186, Feb. 2002.
- [31] D. Raphaeli, "Distribution of noncentral indefinite quadratic forms in complex normal variables," *IEEE Trans. Inf. Theory*, vol. 42, no. 3, pp. 1002–1007, May 1996.
- [32] T. Y. Al-Naffouri and B. Hassibi, "On the distribution of indefinite quadratic forms in Gaussian random variables," in *Proc. Int. Symp. Info. Theory*, Seoul, Jun. 2009, pp. 1744–1748.
- [33] R. Yates, "A framework for uplink power control in cellular radio systems," *IEEE J. Sel. Areas Commun.*, vol. 13, no. 7, pp. 1341–1347, Sep. 1995.
- [34] A. Gründinger, R. Bethenod, M. Joham, M. Reimensberger, and W. Utschick, "Optimal power allocation for the chance-constrained vector broadcast channel and rank-one channel approximation," in *Proc. IEEE Int. Wkshp Signal Process. Adv. Wireless Commun.*, Darmstadt, Germany, Jun. 2013, pp. 31–35.
- [35] E. Björnson, M. Bengtsson, and B. Ottersten, "Optimal multiuser transmit beamforming: A difficult problem with a simple solution structure," *IEEE Signal Process. Mag.*, vol. 31, no. 4, pp. 142–148, Jul. 2014.
- [36] M. Biguesh and A. B. Gershman, "Training-based MIMO channel estimation: A study of estimator tradeoffs and optimal training signals," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 884–893, Mar. 2006.
- [37] G. S. Smith, "A direct derivation of a single-antenna reciprocity relation for the time domain," *IEEE Trans. Antennas Propag.*, vol. 52, no. 6, pp. 1568–1577, 2004.
- [38] F. Kaltenberger, H. Jiang, M. Guillaud, and R. Knopp, "Relative channel reciprocity calibration in MIMO/TDD systems," in *Proc. Future Network and Mobile Summit, 2010*, 2010, pp. 1–10.
- [39] H. V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. Springer-Verlag, 1994.
- [40] E. Song, Q. Shi, M. Sanjabi, R.-Y. Sun, and Z.-Q. Luo, "Robust SINR-constrained MISO downlink beamforming: When is the semidefinite programming relaxation tight?" *EURASIP J. Wireless Commun. Networking*, vol. 2012, no. 243, 2012.
- [41] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multi-antenna multiuser communication—Part I: Channel inversion and regularization," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 195–202, Jan. 2005.
- [42] F. Sohrabi, "Robust power loading for the TDD MISO downlink with outage constraints," Master's thesis, McMaster University, 2013, available online: <http://digitalcommons.mcmaster.ca/opensiderations/8377/>.
- [43] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed. Athena Scientific, 1999.
- [44] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski, *Robust Optimization*. Princeton University Press, 2009.
- [45] R. Stridh, M. Bengtsson, and B. Ottersten, "System evaluation of optimal downlink beamforming with congestion control in wireless communication," *IEEE Trans. Wireless Commun.*, vol. 5, no. 4, pp. 743–751, Apr. 2006.
- [46] F. Sohrabi and T. N. Davidson, "Coordinate update algorithms for robust power loading for the MISO downlink with outage constraints and Gaussian uncertainties," in *Proc. Int. Conf. Acoust., Speech, Signal Process.*, Vancouver, May 2013, pp. 4769–4773.
- [47] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951–963, 2003.
- [48] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, Jr., "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.



processing.

Foad Sohrabi (S'13) received his B.A.Sc. degree in 2011 from the University of Tehran, Tehran, Iran, and his M.A.Sc. degree in 2013 from McMaster University, Hamilton, ON, Canada, both in Electrical and Computer Engineering. Since September 2013, he has been a Ph.D student at University of Toronto, Toronto, ON, Canada. From July to December 2015, he was a research intern at Bell-Labs, Alcatel-Lucent, in Stuttgart, Germany. His main research interests include MIMO communications, optimization theory, wireless communications, and signal



Timothy N. Davidson (M'96, SM'15) received the B.Eng. (Hons. I) degree in electronic engineering from the University of Western Australia (UWA), Perth, in 1991 and the D.Phil. degree in engineering science from the University of Oxford, U.K., in 1995.

He is a Professor in the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada, where he is currently serving as Chair of the Department. Previously, he has served as Acting Director of the School of Computational Engineering and Science for two years, and as Associate Director for three years. His research interests lie in the general areas of communications, signal processing, and control.

Dr. Davidson received the 1991 J. A. Wood Memorial Prize from UWA, the 1991 Rhodes Scholarship for Western Australia, and a 2011 Best Paper Award from the IEEE Signal Processing Society. He has served as an Associate Editor of the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, and the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II. He has also served as a Guest Co-Editor of issues of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS, the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING, and the *EURASIP Journal on Advances in Signal Processing*. He was a General Co-Chair for the 2014 IEEE International Workshop on Signal Processing Advances in Wireless Communications, a Technical Program Co-Chair for the 2014 IEEE Global Conference on Signal and Information Processing, and the Technical Chair for the 2015 Asilomar Conference on Signals, Systems and Computers. Dr. Davidson has also served as the Chair of the IEEE Signal Processing Society's Technical Committee on Signal Processing for Communications and Networking. He is a Registered Professional Engineer in the Province of Ontario.